

Industrial Robotic Equipment: Presence, Exposure, and Capital Expenditures

Beginning with the 2018 ASM, the Census Bureau introduced three new questions to the Annual Survey of Manufactures (ASM) on the presence and use of industrial robots. In response to growing interest and high demand by industry, policy makers, and researchers, these questions fill a data gap by providing the first official federal statistics measuring industrial robotics. These measures allow researchers to understand the relationship between robot adoption and intensity of use, and their impact on employment and productivity. The ASM collects the following information for surveyed manufacturing plants: 1) capital expenditures for industrial robotic equipment, 2) the number of industrial robots present, and 3) the number of industrial robots purchased that year. The new experimental ASM robotics data product includes the following measures:



- National-level estimates of the percent of manufacturing plants with robots, employees exposed to robots, and capital expenditures for robotic equipment.
- Estimates of the percent of manufacturing plants with robots, employees exposed to robots, and capital expenditures for robotic equipment by 3-digit industry, state, and plant size.
- Estimates of the size distribution of manufacturing plants with and without robots.

Capital expenditures for industrial robotic equipment and the number of industrial robots purchased and in use were collected from establishments that are eligible to be sent the ASM form (the mail stratum). For the purpose of this survey, the terms establishment and plant are synonymous. The mail stratum of the Annual Survey of Manufactures (ASM) is a sample survey of approximately 50,000 multi-unit and large single-unit manufacturing establishments. The survey includes manufacturing establishments with one or more paid employees, including establishments that use leased employees. The 2018 ASM sample was based on the sample selected for the 2014 ASM and supplemented annually with births. For the 2019 ASM, a new sample of approximately 50,000 establishments was selected from the universe of manufacturing establishments in the 2017 Economic Census and supplemented with 2018 births. Published data are based only on the mail stratum; no imputation was performed for plants ineligible to be sent a form (the nonmail stratum). The 2018 and 2019 ASM Robotics tabulations are therefore representative of the mail stratum of the ASM only (see [Annual](#)

[Survey of Manufactures Methodology \(census.gov\)](https://www.census.gov/manufacturing/2017naics)). Establishments in the nonmail portion of the ASM account for almost two-thirds of the establishments in the manufacturing universe but account for less than 6% of the value of shipments at the total manufacturing level.

The data are tabulated using the 2017 North American Industry Classification System (NAICS) code assigned to the establishment based on the mail stratum. All estimates are weighted using the ASM sample weights.

The 2018 and 2019 ASM Robotics tabulations use a broad concept of the presence of robots at a plant described in greater detail below. Plants are identified as having robots present or not. Likewise, employees at plants using robots are considered exposed to robots.

Definitions

Robot Presence

Plants are classified as having robots present based on responses to the robotics questions. Plants that report positive values for active robots, capital expenditures on robotic equipment, purchases of new robots, or those who indicate in write-in information they have robots but cannot provide details, are classified as having robots present at the plant. Each of the three robotics questions also includes a "Check if None" checkbox. This provides an additional way for a respondent to report that they have not purchased, nor do they have or use robotics. Plants that select all "Check if None" checkboxes are considered respondents with no robot presence. Respondents that fail to respond to a question or its checkbox and do not report a positive value in at least one of the questions are imputed.

Percent of Plants with Robots

The percentage of plants with robots in a tabulation group is the total number of plants with robots in the group divided by the total number of plants in the group.

Percent of Employees Exposed to Robots

The percentage of employees exposed to robots in a tabulation group is the total number of employees at plants with robots in the group divided by the total number of employees at plants in the group.

Industry Demeaned Percent of Plants with Robots and Employees Exposed to Robots by State

The industry demeaned share of plants with robots is the share of plants with robots in a state after controlling for differences in industry composition. This measure abstracts away from state-level variation in the share of plants with robotics that is simply due to industry composition. A state with more activity in industries that use robotics more intensely will naturally have a higher share of plants with robots. The demeaned measure, on the other hand, measures the *relative* intensity of robot use controlling for national industry-level shares. To compute this measure, we subtract the national industry-level estimates for the share of plants with robots from the state's share for each industry. This yields for each industry the state's deviation in its industry share relative to the national industry share. We take the mean industry deviations for each state and add to it the national share. We average industry deviations for each state. A state's average may be negative if their industry-level shares are lower than the national industry shares. To address this, we add the national total manufacturing share to each state's average.

Capital Expenditures for Industrial Robotic Equipment

Capital expenditures for industrial robotic equipment are the total dollar amount of capital expenditures for new and used industrial robotic equipment, including software, installation, and other one-time costs for each industry or geographic group. Plants that do not respond to this question are imputed. The data are then weighted by the ASM sample weights to get estimates of the totals.

This item is collected as a special inquiry on the ASM form, separate from the standard question asking about the dollar value of capital expenditures on buildings and structures and machinery and equipment (and the further breakout of machinery and equipment into autos, computers, and other machinery). Additionally, the data on robotic equipment capital expenditures are edited, imputed, weighted, and tabulated independently from the capital expenditures data collected and tabulated for the ASM publication tables.

Imputation for Missing Data and Survey Nonresponse

Imputation was used to address missing data either due to item or survey nonresponse. Significant improvements were made to the imputation algorithms between the release of the 2018 and 2019 data years. Previously released estimates for 2018 were revised using the new methodology. Plants with an imputed classification for the presence of robots accounted for approximately 40% of the mailed sample in 2018. A similar percentage of the mailed sample had an imputed value of capital expenditures on robotic equipment in 2018. Plants with an

imputed classification for the presence of robots went up to approximately 45% in the 2019 ASM. The percentage of imputed cases for capital expenditures of robotic equipment also jumped to roughly 45%. The 2019 ASM had a new sample, and the data was collected in the midst of the coronavirus pandemic in calendar year 2020. Separate imputation models were used to impute the presence of robots and capital expenditures on robotic equipment. Numerous regression and machine learning models were considered including logistic propensity scores, random forest, decision trees, and k-neighbor models. Using a 70/30 train/test split, the quality of each imputation model was compared using a Euclidean normalization of the L1 and L2 errors of plant and employee percentages in the resulting state and industry tabulations. The model with the lowest composite error was used in the final imputation of the presence of robotics. Imputed values for capital expenditures on robotic equipment are not used to identify the presence of robots.

Researchers with approved access to the restricted use microdata can read more about the imputation methods in (CES Technical Note Reference). Goldschlag, Jones, Miranda, and Smith (2022) find that simple linear and logistic regressions models with state, industry, size, and age perform better than the more complex machine learning models with and without feature selection models. The results also suggest that the establishment-level accuracy of imputed values is poor. Even within detailed industry and state cells it is very difficult to identify plants exposed to robots. Despite this, state and industry-level tabular estimates are robust.

Disclosure Avoidance

Disclosure is the release of data that reveals information or permits deduction of information about a particular survey unit through the release of either tables or microdata. Disclosure avoidance is the process used to protect each survey unit's identity and data from disclosure. Using disclosure avoidance procedures, the Census Bureau modifies or removes the characteristics that put information at risk of disclosure. Although it may appear that a table shows information about a specific survey unit, the Census Bureau has taken steps to disguise or suppress a unit's data that may be "at risk" of disclosure while making sure the results are still useful.

Cell suppression is a disclosure avoidance technique that protects the confidentiality of individual survey units by withholding cell values from release and replacing the cell value with a symbol, usually a "D". If the suppressed cell value were known, it would allow one to estimate an individual survey unit's response too closely.

The process of suppression does not usually change the higher-level totals. Values for cells that are not suppressed remain unchanged.

The Census Bureau has reviewed the data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied. (Approval IDs: (CBDRB-FY21-ESMD001-008, CBDRB-FY22-251)